

ABSTRACT OF THE DISCLOSURE

A method of expressing a target symbol sequence T relative to a reference symbol sequence R including identifying a first LCS of symbols in sequences T and R, defining the first LCS as a root node of a tree, the root node being a parent node, for each portion of the sequence T that precedes or succeeds the LCS in the sequence T, where there is a portion of the sequence R corresponding to the portion of the sequence T, identifying a subsequent LCS of symbols in the portions, if the subsequent LCS is identified, defining the subsequent LCS as a child node of the parent node, if the subsequent LCS is not identified, defining a child leaf of the parent node, and, where there is no portion of the sequence R corresponding to the portion of the sequence T, defining a child leaf of the parent node.